**Note on the Optimization of priors**

**Aim**

When inverting several models for further model comparison, the choice of priors is important.

Different choice of priors can lead to different results. Here we propose to optimize priors so to minimize model selection errors on simulated data.

**Theory**

We consider different models (dealing with partition is an easy extension)

For any data sampled by a model , we define a (selection) error as the posterior probability of models that did not generate the data

For a given model generating data, we would like to compute the expected error

This depends on priors we choose to perform the inversion.

Different priors lead to different posterior on models, and different predictive density.

Here, the best way to minimize the expected error is to choose priors so that predictive distributions of the different models have as little overlap as possible. (not very interesting)

Taken all models together, we would like to compute the expected error marginalized on models

With being our prior knowledge on models.

Here again, error is minimized by having as little overlap between the predictive densities. (not interesting)

**Issue**

This has no interest. Hence, the optimization leads to non overlapping predictive densities.

The optimized priors are priors leading to deterministic outputs.

The resulting priors might also have be irrelevant for the kind of processes we aim to model.

The predictive densities might also be far away from the kind of data we will in the end observe

**Solution 1**

We propose to approximate this expected cost, on data sampled from priors leading to data close to what is actually observed.

We approximate expected error as follows

The trade-off here is that to minimize expected error, priors must be more predictive of the data they generated than other models (i.e. low ). This means that whatever the data, the model that generated it should explain best

Previous problem is then solved.

This is what we try to implement

**Documentation of the code : VBA\_optimPriors**

* Data are simulated
* For each model, all simulations are inverted for all the possible priors
* Model posterior is computed for all configurations for all priors
* Error is computed on all these priors